

DOI:10.19651/j.cnki.cmt.2107593

基于 MPE 与 PSO-SVM 的滚动轴承故障诊断<sup>\*</sup>刁宁昆<sup>1,2</sup> 马怀祥<sup>1,2</sup> 王金师<sup>2</sup> 刘 帅<sup>2</sup>

(1. 石家庄铁道大学 河北省大型工程机械装备制造协同创新中心 石家庄 050043;

2. 石家庄铁道大学 机械工程学院 石家庄 050043)

**摘要:** 滚动轴承是旋转机械的重要部件之一,针对滚动轴承故障诊断问题,提出了一种多尺度排列熵(MPE)与粒子群优化(PSO)的支持向量机(SVM)相结合的算法。利用 MPE 方法得到轴承故障信号的故障特征,并将其作为特征向量输入 PSO-SVM 模型中,使用凯斯西储大学轴承故障数据进行验证,发现该方法可以有效进行滚动轴承的故障识别。同时将该方法与多尺度排列熵结合传统的 SVM 方法以及使用网格搜索优化的 SVM 方法所得故障分类结果进行比较,发现该方法在滚动轴承故障诊断的时效性以及准确率方面具有一定的优越性。

**关键词:** 多尺度排列熵;粒子群优化;支持向量机;滚动轴承

**中图分类号:** TH133.33+1;TP181 **文献标识码:** A **国家标准学科分类代码:** 460

## Fault diagnosis of rolling bearing based on MPE and PSO-SVM

Diao Ningkun<sup>1,2</sup> Ma Huaixiang<sup>1,2</sup> Wang Jinshi<sup>2</sup> Liu Shuai<sup>2</sup>

(1. Hebei Provincial Collaborative Innovation Center of Large Construction Machinery Manufacturing, Shijiazhuang Tiedao University, Shijiazhuang 050043, China; 2. School of Mechanical Engineering, Shijiazhuang Tiedao University, Shijiazhuang 050043, China)

**Abstract:** Rolling bearing is one of the important parts of rotating machine. Aiming at the problem of rolling bearing fault diagnosis, this paper proposes an algorithm combining multiscale permutation entropy (MPE) and support vector machine (SVM) optimized by particle swarm optimization (PSO). The fault characteristics of the bearing fault data was obtained by the MPE method, fitting as a feature vector into the PSO-SVM model, using Case Western Reserve University bearing dataset for verification. It is found that this method can effectively identify the fault of the rolling bearing. This method is compared with the fault classification results obtained by combining the multi-scale permutation entropy with the traditional SVM method and the SVM method optimized by grid search. It is found that the method proposed has certain advantages in the efficiency and accuracy of rolling bearing fault diagnosis.

**Keywords:** multiscale permutation entropy; particle swarm optimization; support vector machine; rolling bearings

## 0 引 言

滚动轴承在现代工业中扮演着重要角色,其故障与否影响着机械设备整体的运行<sup>[1]</sup>。常用的滚动轴承故障诊断方法包括时域和频域的分析。时域信号只能观察到任意时刻的幅值大小,傅里叶可以分析信号频率成分及其对应幅值,但其得到信号频谱需要利用信号的全部时域信息,且无法反映信号频率成分与时间的关系。小波变换运用局部化的思想,同时可根据频率高低自动调整窗口大小,具有很好的自适应能力<sup>[2]</sup>。峭度指标可以直接反映轴承所受冲击,但是其鲁棒性较弱,不适用于瞬态信号的监测<sup>[3]</sup>。Huang 提出的经验模态分解方法(empirical mode decomposition,

EMD)<sup>[4]</sup>将信号根据自身尺度特征进行分解,分解得到的每一个本征模函数都代表了原始信号中存在的一种内涵模态分量,包含了原信号在不同时间尺度上的局部特征信号,但是存在模态混淆和端点效应<sup>[5]</sup>等一系列弊端。从机械振动的非平稳、非线性角度出发,上述方法难免存在一定局限性。

排列熵(permutation entropy, PE)<sup>[6]</sup>是由 Bandt 等提出的一种检测一维时间序列随机性和动力学突变的方法,相比于一般的线性分析方法,具有计算简单快速、鲁棒性强等优点。排列熵的概念提出后,在生物组织识别<sup>[7]</sup>、风电功率预测<sup>[8]</sup>等方面都得到了一定的应用。Yan 等将排列熵<sup>[9]</sup>

收稿日期:2021-08-15

<sup>\*</sup> 基金项目:国家自然科学基金(11872254)、中铁十四局集团有限公司芜湖长江隧道建设指挥部工程科研项目(ZTSSJ-WHSD-GCKY-2021-002)资助

应用于旋转机械的状态监测,程军圣等<sup>[10]</sup>将排列熵与基于变量预测的模式识别(variable predictive model based class discrimination, VPMCD)相结合进行滚动轴承故障诊断。排列熵能够对振动信号的动态变化进行有效检测和放大,且可以表征滚动轴承的工作状态。但是,排列熵类似于一般的单尺度非线性参数,仅反映时间序列在单一尺度上的随机性与动力学突变<sup>[11]</sup>。Aziz在排列熵的基础上提出的多尺度排列熵<sup>[12]</sup>的概念,在多尺度上衡量时间序列随机性,并将其与多尺度熵方法进行对比,得出了多尺度排列熵具有更好鲁棒性的结论。

本文将多尺度排列熵(multiscale permutation entropy, MPE)与粒子群优化(particle swarm optimization, PSO)的支持向量机(support vector machine, SVM)方法相结合进行滚动轴承故障诊断。当滚动轴承发生故障时,其振动信号随机性及动力学行为都会产生突变。使用多尺度排列熵量化这种突变行为,并将故障信号熵值作为特征向量输入至粒子群优化的支持向量机中,用以分辨不同类型的轴承故障。使用实验数据进行验证,最终表明,本文提出的MPE以及PSO-SVM结合方法在滚动轴承故障诊断的时效性和准确率上具有一定优势。

## 1 多尺度排列熵

### 1.1 排列熵算法

排列熵算法可以描述一维时间信号的随机性,设一长度为 $n$ 的时间序列 $\{x(i), i = 1, 2, \dots, n\}$ ,对其进行相空间重构,可得如下矩阵:

$$\begin{bmatrix} x(1) & x(1+\tau) & \dots & x(1+(m-1)\tau) \\ \vdots & \vdots & \vdots & \vdots \\ x(i) & x(i+\tau) & \dots & x(i+(m-1)\tau) \\ x(K) & x(K+\tau) & \dots & x(K+(m-1)\tau) \end{bmatrix}, i = 1, 2, \dots, K \quad (1)$$

其中, $m$ 为嵌入维度; $\tau$ 为时间延迟; $K = n - (m-1)\tau$ 。重构矩阵的每一行看为一独立分量,共 $K$ 个重构分量,将第 $i$ 个分量 $(x(i), x(i+\tau), \dots, x(i+(m-1)\tau))$ 中的每一个因素都按数值由小到大重新排列,即:

$$x[i+(j_1-\tau)] \leq x[i+(j_2-\tau)] \leq \dots \leq x[i+(j_m-\tau)] \quad (2)$$

此时,若存在相邻值相等的情况,则按照 $j$ 值大小进行排序,即当 $j_a < j_{a-1}$ 时,有:

$$x[i+(j_a-\tau)] \leq x[i+(j_{a-1}-\tau)] \quad (3)$$

因此,将任意时间序列进行重构后均可得到一组符号序列:

$$S(l) = \{j_1, j_2, \dots, j_m\} \quad (4)$$

其中, $l = 1, 2, \dots, k, k \leq m!$ , $m$ 维相空间映射不同的符号序列 $j_1, j_2, \dots, j_m$ ,共计 $m!$ 种, $S(l)$ 为其中一种。

对每种符号出现的概率 $P_1, P_2, \dots, P_k$ 进行计算,以Shannon熵的形式,此时间序列 $\{x(i), i = 1, 2, \dots, n\}$ 的排

列熵可以定义为:

$$m = - \sum_{j=1}^k P_j \ln P_j \quad (5)$$

当 $P_x = 1/m!$ 时, $H_p(m)$ 达到最大值 $\ln(m!)$ ,对 $H_p(m)$ 进行归一化处理,即:

$$H_p = H_p(m) / \ln(m!) \quad (6)$$

此时显然 $H_p$ 的取值范围为 $0 \leq H_p \leq 1$ , $H_p$ 的值反映了时间序列的随机程度,并与其呈正相关,同时也反映并放大了时间序列的细微变化。

### 1.2 排列熵参数选取

在进行排列熵计算之前,需考虑信号长度 $N$ 、嵌入维数 $m$ 和时间延迟 $\tau$ 的变化对结果的影响。若 $m$ 取值过小,重构向量中很难包含进行下一步分析所需的足够信息,难以有效反映信号随机性;若 $m$ 取值过大,此时间序列会因相空间重构而均匀化,不仅计算相当耗时,同时也无法体现时间序列的细微变化。

首先分析时间延迟 $\tau$ 在排列熵计算过程中的影响。取一段长度为1024的高斯白噪声信号,图1所示为不同时延 $\tau$ 下排列熵值随嵌入维数的变化关系。

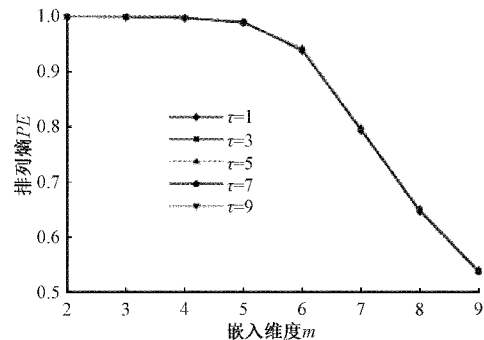


图1 高斯白噪声信号在不同时延下的排列熵值

显然,时间延迟 $\tau$ 的大小对信号熵值的计算并不会产生很大的影响,本文取 $\tau = 1$ 。

接下来考虑信号长度对信号排列熵的影响,取长度分别为128、256、512、1024、2048的高斯白噪声,记为 $PE_1 \sim PE_5$ ,图2所示为 $PE_1 \sim PE_5$ 在不同嵌入维度下的排列熵值,表1所示为它们在不同嵌入维度下的差值。

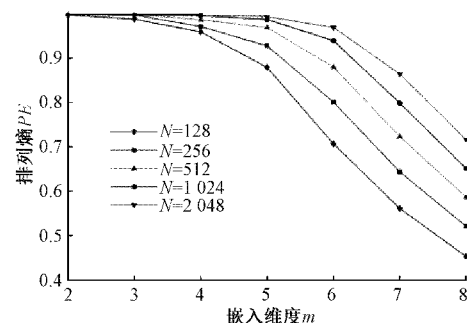


图2 不同长度信号的排列熵

表 1 不同长度信号在不同嵌入维度下的排列熵差值

$m$	2	3	4	5	6	7	8
$PE_5 \sim PE_4$	0.000 1	-0.000 1	0.002 4	0.005 9	0.038 1	0.069 1	0.064 3
$PE_1 \sim PE_3$	0	0.000 8	0.001 4	0.010 5	0.057 0	0.077 3	0.065 6
$PE_3 \sim PE_2$	0.000 4	0.001 5	0.005 7	0.030 3	0.071 7	0.079 0	0.067 2
$PE_2 \sim PE_1$	0.000 6	0.000 2	0.015 6	0.058 1	0.089 3	0.080 7	0.067 6

可以看出,嵌入维度越大,则对信号长度的要求就越大。举例而言,当嵌入维度取 6 时,长度为 1 024 的高斯白噪声信号的排列熵值与长度为 512 的高斯白噪声信号的排列熵值的差为 0.057 0,而长度为 2 048 与长度为 1 024 的高斯白噪声信号的排列熵的差值为 0.038 1,明显小于前一个值。当嵌入维度的值取 5 时,长度为 1 024 和 512 的信号熵值的差为 0.010 5,而长度为 2 048 和 1 024 的信号熵值之差仅为 0.005 9。本文中,取数据长度为 2 048,基于重构向量中包含的信息量以及计算时间两点考虑,取嵌入维度为 6。

## 2 PSO-SVM 方法

### 2.1 支持向量机

支持向量机是 Vapnik 等提出的一种机器学习算法<sup>[13]</sup>,一般用于小样本的分类预测,在解决复杂非线性问题上具有很广泛的应用。SVM 方法面对线性不可分情况时引入核映射方法,将其转化为高维空间的线性可分问题从而实现分类。核函数的选择是 SVM 的关键问题,常用的核函数有 Gaussian 核函数(radial basis function, RBF)、多项式函数、S 形函数等<sup>[14]</sup>。RBF 核函数中仅需确定一个参数  $\sigma$ ,本文采用 RBF 核函数。此外,惩罚参数  $C$  在 SVM 分类过程中也会产生一定程度的影响,因此,需引入一种对参数  $\sigma$  和  $C$  的寻优方法以提高分类精度。

### 2.2 粒子群优化算法

粒子群优化算法<sup>[15]</sup>是由 Kennedy 和 Eberhart 提出的一种基于群智能的优化算法,通过个体间的协作与竞争来进行复杂解空间的寻优,不仅收敛速度快,也没有过多参数需要设置,因此对于 SVM 的参数优化具有很好的适用性。粒子群优化算法的原理为:在空间中定义一群粒子,粒子状态可以用初始速度、位置及适应度值进行描述,粒子可以记忆在设计空间中搜寻到的最佳位置,其适应度值根据适应度函数计算得出。以粒子的速度为依据判断粒子飞行的方向和距离,粒子可在最优解空间中飞行搜索。在不断迭代寻优的过程中,粒子通过比较适应度以及个体、全局极值来进行其速度与位置的更新,其更新公式为:

$$v_i(t+1) = \omega v_i(t) + c_1 R_1 [R_g^b(t) - x_i(t)] + c_2 R_2 [R_p^b(t) - x_i(t)] \quad (7)$$

$$x_i(t+1) = x_i(t) + \varphi v_i(t+1) \quad (8)$$

其中,  $t$  代表迭代次数;  $v_i(t)$  为第  $i$  个粒子的第  $t$  次迭代速度;  $\omega$  为惯性权重;  $c_1, c_2$  为认知系数,分别影响局

部搜索和全局搜索能力;  $R_1, R_2$  为均匀分布随机数;  $R_g^b(t)$  为第  $i$  个粒子的历史最优位置;  $R_p^b(t)$  为群体历史最优位置;  $x_i(t)$  为粒子在第  $t$  次迭代的位置;  $\varphi$  为收缩因子,其作用为控制粒子速度,使其在特定范围内。使用 PSO 算法优化 SVM 参数的流程如图 3 所示。

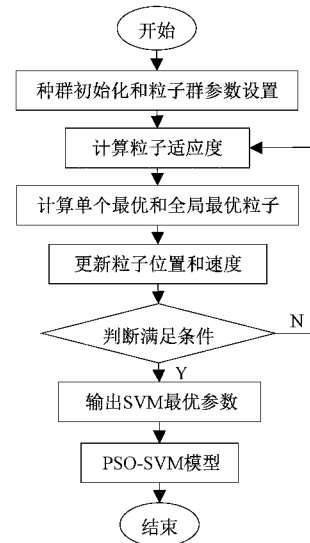


图 3 PSO 算法优化 SVM 参数的流程

## 3 基于 MPE 和 PSO-SVM 的滚动轴承故障诊断方法

为验证上述多尺度排列熵结合粒子群优化支持向量机方法的有效性,使用凯斯西储大学滚动轴承数据对该方法的实用性进行进一步的验证。实验选取轴承型号为 6205-2RS JEM SKF 型深沟球轴承,采样频率为 12 kHz,电机转速为 1 797 r/min,负载为 0 HP。轴承损伤为电火花加工的单点损伤,故障直径为 0.177 8 mm,深度为 0.279 4 mm。轴承共 4 种状态类型,分别为正常状态、内圈故障、滚动体故障和外圈故障。各状态类型所对应的时域波形如图 4 所示。

取每组数据长度为 2 048,嵌入维度为 6,时间延迟为 1,最大尺度因子为 15。4 种状态的数据各截取一段,求取其排列熵值,不同状态下排列熵随尺度因子变化的关系如图 5 所示。

可以看出,不同状态类型的滚动轴承对应的排列熵值大小不同。当尺度因子取 1 时,正常状态下的振动信号排列熵值和故障状态下的排列熵值有较大差距,但是 3 种故

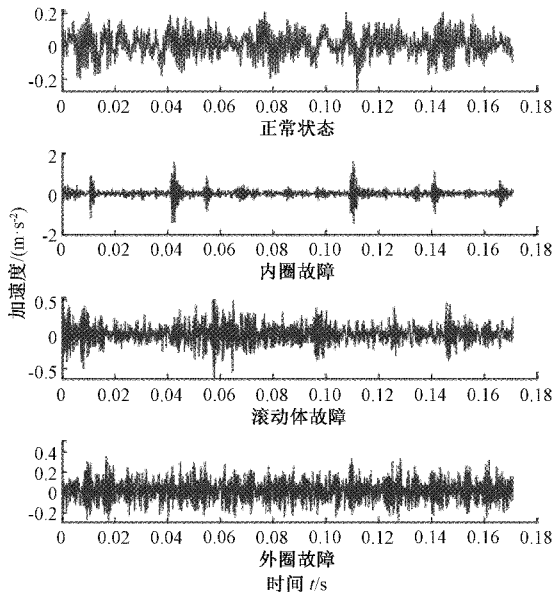


图 4 4 种状态轴承的时域波形图

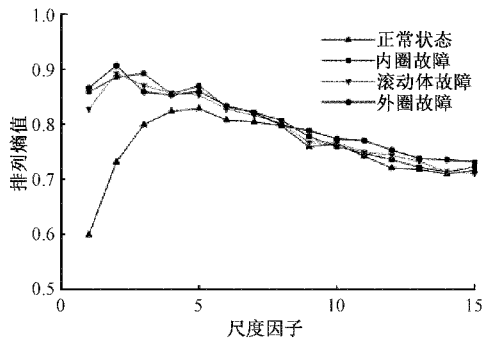


图 5 4 种状态下的多尺度排列熵值

障状态下的振动信号排列熵值却比较接近,因此很难直观

判断故障的类型。且当尺度因子逐渐增大时,4 种状态信号的排列熵值逐渐趋向接近。因此,仅在前几个尺度的排列熵值中包含较多信息。若取过多尺度的熵值,不仅造成信息的冗余,同时也会降低分类器的运算速度。因此,本文中取前 5 个尺度的排列熵组成特征向量。

对上述各种状态类型的数据各选取样本数为 60,数据长度为 2 048 的样本共 4 组,总计 240 个样本。将每种状态 60 个样本中的前 30 个作为训练样本,后 30 个作为测试样本。计算出各样本前 5 个尺度的排列熵,作为特征向量。将每种状态的前 30 个样本的特征向量输入到 PSO-SVM 分类器中,再将后 30 个样本输入到已经训练好的分类器中进行识别。所得分类结果如图 6 所示。

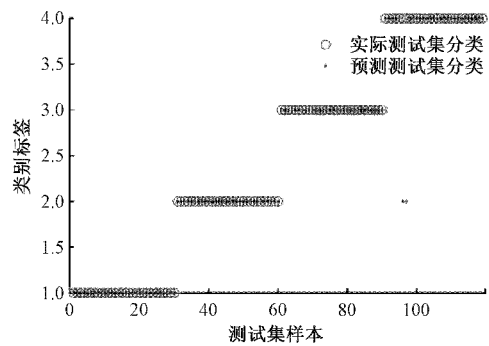


图 6 PSO-SVM 分类结果

在 120 个测试样本中,仅两个样本被错分,准确率达到了 98.33%,证明本文提出的方法在滚动轴承的故障诊断方面具有很好的效果。为了突出 MPE 与 PSO-SVM 相结合进行故障诊断的优势,将其他两种方法与其进行对比,分别是 MPE 结合 SVM,以及 MPE 结合使用网格搜索 (grid search, GS)<sup>[16]</sup> 优化的 SVM,所得到的分类准确率结果如表 2 所示。

表 2 不同方法进行故障诊断的准确率

诊断方法	正确识别样本数				总准确率/%
	正常状态	内圈故障	滚动体故障	外圈故障	
MPE 和 PSO-SVM	30	30	30	28	98.33
MPE 和 SVM	30	29	18	28	89.17
MPE 和 GS 优化 SVM	30	29	30	28	97.50

可以看出,3 种方法在识别外圈故障时均出现了两个错误识别,但是本文提出的方法在识别其他 3 种状态时完全准确,而另外两种方法均未达到这种准确率,说明将 MPE 与 PSO-SVM 结合的方法具有一定优越性,且 3 种方法的计算耗时都比较接近。文献[17]中,将排列熵的结果分别输入到 BP 神经网络和支持向量机中,所得结果显示, BP 神经网络分类器不仅分类准确率要低于支持向量机,而且在训练时间上也较长。

#### 4 结 论

本文将多尺度排列熵与粒子群优化的支持向量机方法结合进行滚动轴承的故障诊断。多尺度排列熵可以用来检测时间序列的动力学突变,对滚动轴承早期和晚期的变化表现得比较敏感。而粒子群算法和支持向量机的结合,可以充分利用 SVM 在处理小样本回归问题上具有的独特优越性及 PSO 全局优化的特点。将该方法应用于实

验数据的分析,达到了98.33%的准确率,相对于仅使用SVM或GS优化参数的SVM来说,在准确率方面具有一定优越性,同时该方法与BP神经网络相比,计算耗时也相对较短。因此,该方法在滚动轴承的故障诊断以及状态分析方面具有很好的适用性。

### 参考文献

- [1] 吴涛,熊新,吴建德,等.基于QH-ITD和AMCKD的滚动轴承故障诊断研究[J].电子测量与仪器学报,2020,34(4):79-89.
- [2] 陈乐鹏,谭晓东. ip-iq 的改进算法与小波变换谐波检测方法研究[J]. 电子测量技术,2020,43(19):69-74.
- [3] 张龙,毛志德,杨世锡,等.基于包络谱带峭度的改进谱峭度方法及在轴承诊断中的应用[J]. 振动与冲击,2018,37(23):171-179.
- [4] 马云飞,贾希胜,胡起伟,等.基于EMD和BCS的振动信号数据修复方法[J]. 仪器仪表学报,2019,40(3):154-162.
- [5] 范勇,王鹏,张宁超.基于改进EMD与SOM神经网络的自动机驱动机构故障诊断研究[J]. 国外电子测量技术,2020,39(8):5-10.
- [6] 熊国良,甄灿壮,张龙,等.基于CEEMDAN多尺度排列熵的轴承故障智能识别 Fisher-GG 聚类方法[J]. 噪声与振动控制,2020,40(6):1-7,28.
- [7] 刘备,胡伟鹏,邹孝,等.基于变分模态分解与多尺度排列熵的生物组织变性识别[J]. 物理学报,2019,68(2):259-267.
- [8] 曹威,刘长良,王梓齐,等.基于MEEMD与排列熵的风电功率超短期预测研究[J]. 可再生能源,2019,37(3):439-444.
- [9] GAO L. Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines[J]. Mechanical Systems and Signal Processing, 2012, 29(5): 474-484.
- [10] 程军圣,马兴伟,杨宇.基于排列熵和VPMCD的滚动轴承故障诊断方法[J]. 振动与冲击,2014,33(11):119-123.
- [11] 郑近德,程军圣,杨宇.基于LCD和排列熵的滚动轴承故障诊断[J]. 振动. 测试与诊断,2014,34(5):802-806,971.
- [12] 赵荣珍,李宪蒲,邓林峰. EWT 多尺度排列熵与 GG 聚类的轴承故障辨识方法[J]. 振动. 测试与诊断,2019,39(2):416-423,451.
- [13] 周绍磊,廖剑,史贤俊. RBF-SVM 的核参数选择方法及其在故障诊断中的应用[J]. 电子测量与仪器学报,2014,28(3):240-246.
- [14] 李莎,潘宏侠,都衡.基于EEMD信息熵和PSO-SVM的自动机故障诊断[J]. 机械设计与研究,2014,30(6):26-29,33.
- [15] 李益兵,王磊,江丽.基于PSO改进深度置信网络的滚动轴承故障诊断[J]. 振动与冲击,2020,39(5):89-96.
- [16] 郑近德,程军圣,杨宇.多尺度排列熵及其在滚动轴承故障诊断中的应用[J]. 中国机械工程,2013,24(19):2641-2646.
- [17] 吴峰,陈后金,姚畅,等.基于网格搜索的PCA-SVM道路交通标志识别[J]. 铁道学报,2014,36(11):60-64.

### 作者简介

刁宁昆,硕士研究生,主要研究方向为机电设备的状态监测与故障诊断。

E-mail:1591312122@qq.com

马怀祥(通信作者),教授,主要研究方向为机电设备的状态监测与故障诊断。

E-mail:416826624@qq.com